

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>7</sup> : <b>C07H 21/04, A01H 5/00, C12N 15/09, 15/14, 15/00</b>		A1	(11) International Publication Number: <b>WO 00/49035</b> (43) International Publication Date: <b>24 August 2000 (24.08.00)</b>
(21) International Application Number: <b>PCT/US00/04287</b>		(81) Designated States: AB, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, I.S., MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: <b>18 February 2000 (18.02.00)</b>		Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(30) Priority Data: <b>60/120,682 19 February 1999 (19.02.99) US</b>			
(71) Applicant: <b>THE GENERAL HOSPITAL CORPORATION (US/US); 55 Fruit Street, Boston, MA 02114 (US).</b>			
(72) Inventor: <b>SHEEN, Jen; 9 Hawthorne Place, Apartment SJ, Boston, MA 02140 (US).</b>			
(74) Agent: <b>ELBING, Karen, L.; Clark &amp; Elbing, LLP, 176 Federal Street, Boston, MA 02110 (US).</b>			
(54) Title: <b>GENE SILENCING</b>			
(57) Abstract <p>Disclosed is a method for silencing the expression of an endogenous gene in a cell, the method involving overexpressing in the cell a nucleic acid molecule of the endogenous gene and an antisense molecule including a nucleic acid molecule complementary to the nucleic acid molecule of the endogenous gene, wherein the overexpression of the nucleic acid molecule of the endogenous gene and the antisense molecule in the cell silences the expression of the endogenous gene.</p>			

***FOR THE PURPOSES OF INFORMATION ONLY***

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BK	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MIN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KC	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DK	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EM	Estonia						

## GENE SILENCING

5

### Background of the Invention

This invention relates generally to the silencing of gene expression, as well as to expression vectors useful for silencing the expression of a target gene.

Cells manufacture proteins by transcribing genes encoded by DNA. In particular, one of the DNA strands of the gene is transcribed by an enzyme, RNA polymerase, to produce mRNA. The mRNA molecule has a base sequence that is complementary to that of the transcribed DNA strand. The mRNA is then processed by the removal of introns. The base sequence of the mRNA is next translated into the amino acid sequence of a protein molecule by means of the genetic code. This translation process requires many enzymes and a set of transfer RNA molecules, which align the amino acids according to the codon sequence. The translation of mRNA into protein occurs on ribosomes. The translated protein is often referred to as a gene product.

Developmental and physiological processes of cells and organisms require that the gene products necessary to carry out these processes are available in the appropriate amounts and at the appropriate times. For example, certain gene products must be present in all cells for many fundamental physiological processes to occur. Other gene products have tissue- or cell-specificity and are only necessary in certain cells or tissues. Some gene products are continuously present either in all cells or in certain cells or tissues. Still other gene products are required at different times during development.

Gene regulation, in both animals and plants, alters the quantity or quality or both of a gene product. These alterations can be used to ascertain the molecular activities of the normal gene product counterpart. Additionally, gene regulation can be used to manipulate cells and organisms at the genetic level. Gene regulation provides scientists with an expanded ability to study and treat disease processes. In particular, gene regulation techniques have proven especially useful in the elucidation

-2-

and diagnosis of many diseases and abnormalities. Indeed, it is now possible to use gene regulation for therapeutic intervention and treatment at the genetic level. In addition, gene regulation can be used to identify and characterize genes involved in fundamental cellular and developmental processes of animals and plants. The 5 identification and characterization of these genes has previously been hampered by the fact that mutations in such genes are often lethal or are recessive in diploid organisms. Accordingly, the selective inactivation or regulation of genes has many potential uses.

Certain genes are regulated at the level of transcription. Transcriptional 10 regulation is carried out either negatively (repressors) or positively (activators) by a protein factor. Specific protein factors regulate translation of specific mRNAs. It has also become evident that RNAs are involved in regulating the expression of specific genes.

Different approaches have attempted to regulate gene expression by 15 selectively inactivating genes. For example, one approach to gene regulation is gene disruption. Gene disruption is accomplished using recombinant DNA techniques. It is generally a process of sequential elimination of the alleles for a particular gene. The alleles are eliminated by introducing a mutation into the gene, usually by homologous recombination, at the single cell stage of the organism, which renders the 20 gene nonfunctional. Gene disruption is technically difficult and labor intensive. In addition, additional genes are sometimes unintentionally disrupted.

Another gene regulation approach involves the use of antisense RNA. Antisense RNA inhibition of specific mRNAs involves the use of DNA constructs 25 that direct the transcription of an antisense RNA strand. For example, antisense DNA constructs have been prepared by flipping a gene fragment of interest and inserting this sequence between a promoter and a polyadenylation site in the inverse orientation. The RNA transcript obtained from this DNA molecule has a sequence complementary to a target mRNA. The antisense RNA anneals to the mRNA and disrupts normal processing or translation or both. Antisense constructs can be

-3-

introduced into eukaryotic cells by standard transfection, transduction, or microinjection methods and function in both transient and stable transformation assays. Antisense transcripts complementary to the target gene mRNA specifically suppress gene activity.

5       Despite the availability of these and other methods for regulating gene expression, there exists a need for gene regulation methods that are capable of silencing or suppressing a specific gene product or multiple gene products.

#### Summary of the Invention

In one aspect, the invention features a method for silencing the expression 10 of a targeted gene (e.g., an endogenous gene) in a cell. The method generally involves overexpressing in the cell (e.g., an animal cell such as a mammalian cell, a cancer cell, or a plant cell) an isolated nucleic acid molecule of an endogenous gene and an antisense molecule including a nucleic acid molecule complementary to the nucleic acid molecule of the endogenous gene, wherein the overexpression of the 15 nucleic acid molecule of the endogenous gene and the antisense molecule in the cell silences the expression of the endogenous gene. Exemplary genes targeted for silencing include, without limitation, transcription regulatory factors, virally encoded proteins (e.g., human papilloma virus E6 and human immunodeficiency virus tat), structural proteins, metabolic or enzymatic proteins, cytokines, oncogenes, growth 20 factors (e.g., interleukins), gamma interferon, tumor necrosis factor, and granulocyte-macrophage-colony stimulating factor.

In preferred embodiments, the isolated nucleic acid molecule of an endogenous gene generates an untranslatable RNA molecule or a translatable RNA molecule that encodes a non-functional product. In other preferred embodiments, the 25 isolated nucleic acid molecule includes the coding region of the endogenous gene or a portion thereof. While in still other embodiments, the antisense molecule is complementary to a portion of the gene targeted for silencing.

Preferably, the isolated nucleic acid molecule of the endogenous gene used

-4-

in the invention includes an exon (e.g., an exon that is between 20-2,000 base pairs). In preferred embodiments, overexpression of the exon and the antisense molecule silences the expression of a gene family. In other preferred embodiments, overexpression of the exon and the antisense molecule silences the expression of a specific member of the gene family.

In yet other preferred embodiments, the expression of the isolated nucleic acid molecule, antisense molecule, or exon is controlled by an inducible expression control region, a tissue- or cell-specific expression control region, or by a (fully or partially) constitutive expression control region.

In preferred embodiments, the endogenous gene selected for gene silencing encodes a protein, for example, a metabolic enzyme, a structural protein, or a gene product that is associated with a disorder (e.g., an autoimmune disease, cancer, tissue inflammation, or a dysfunction in a metabolic pathway). In still other preferred embodiments, the endogenous gene encodes a gene product that when silenced confers on a plant an agronomically important trait (e.g., fruit ripening, senescence, male sterility, wounding response, disease resistance, or dessication tolerance).

In another aspect, the invention features a method for silencing the expression of an endogenous gene in a cell. The method generally involves expressing in the cell a vector including (i) an expression control region functional in the cell; and (ii) an operably linked DNA molecule including a proximal region and a distal region, the proximal region having substantial sequence identity to an endogenous gene of the cell targeted for silencing, wherein transcription of the DNA generates an RNA molecule having a genetically-engineered dsRNA stem-loop structure based on complementarity between nucleotides found in the distal region of the RNA molecule. In preferred embodiments, the expression vector is overexpressed in the cell.

In still another aspect, the invention features a gene silencing expression vector including (i) an expression control region functional in the cell; and (ii) an

-5-

operably linked DNA molecule including a proximal region and a distal region, the proximal region having substantial sequence identity to an endogenous gene of the cell targeted for silencing, wherein transcription of the DNA generates an RNA molecule having a genetically-engineered dsRNA stem-loop structure based on

5 complementarity between nucleotides found in the distal region of the RNA molecule.

In another aspect, the invention features a method for silencing the expression of an endogenous gene in a cell. The method generally involves expressing in the cell a vector including: (i) an expression control region functional in the cell; and (ii) an operably linked DNA molecule having substantial sequence identity to an endogenous gene of the cell; wherein transcription of the DNA generates an RNA molecule that forms dsRNA structure based on complementarity between nucleotides found in the 5' and 3' regions of the RNA molecule.

In yet another aspect, the invention features a gene silencing expression vector that includes: (i) an expression control region functional in the cell; and (ii) an operably linked DNA molecule having substantial sequence identity to an endogenous gene of the cell; wherein transcription of the DNA generates an RNA molecule that forms a dsRNA structure based on complementarity between nucleotides found in the 5' and 3' regions of the RNA molecule.

In still another aspect, the invention features transgenic plants and non-human transgenic animals that overexpress in a cell (e.g., a non-human animal cell such as a mammalian cell, a cancer cell, or a plant cell) an isolated nucleic acid molecule of an endogenous gene and an antisense molecule including a nucleic acid molecule complementary to the nucleic acid molecule of the endogenous gene. The invention also features transgenic plants and non-human transgenic animals including the expression vector of the invention.

As used herein, by "polypeptide" is meant any chain of amino acids, regardless of length or post-translational modification (for example, glycosylation or phosphorylation).

By "substantially identical" is meant a nucleic acid molecule exhibiting at

-6-

least 40%, preferably 50%, more preferably 80%, and most preferably 90%, or even 95% contiguous sequence identity to a reference sequence. The length of comparison sequences will generally be at least 15-20 nucleotides, preferably at least 50 nucleotides, more preferably at least 75 nucleotides, and most preferably 100

5 nucleotides or more. Sequence identity is measured, for example, using standard sequence analysis software (for example, Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705, Blast, BlastN, BlastP, BlastX, FastA, or PILEUP/PRETTYBOX programs) set to standard parameters.

10 By "nucleic acid" or "nucleic acid molecule" is meant a polymer of deoxyribonucleotides or ribonucleotides, in the form of a separate fragment or as a component of a larger expression vector. For example, nucleic acids can be assembled from cDNA fragments or from oligonucleotides to generate a synthetic gene which is capable of being expressed in a recombinant transcriptional unit, 15 generating transcripts that accumulate stably. Polynucleotide or nucleic acid molecules of the invention therefore include DNA, RNA, cDNA, or synthetic nucleic acid sequences.

By "isolated nucleic acid molecule" is meant a nucleic acid molecule (e.g., 20 DNA) that is free of the genes which, in the naturally-occurring genome of the organism from which the nucleic acid molecule is derived, flank the gene. The term therefore includes, for example, a gene or fragment thereof that is incorporated into a vector; into an autonomously replicating plasmid or virus; or into the genomic DNA of a prokaryote or eukaryote; or that exists as a separate molecule (for example, a cDNA or a genomic or cDNA fragment produced by PCR or restriction endonuclease 25 digestion or a synthetic nucleic acid molecule) independent of other sequences. It also includes a recombinant nucleic acid molecule which is part of a hybrid gene encoding additional polypeptide sequence.

By "antisense molecule" is meant an RNA molecule that includes a nucleotide sequence that is complementary to a messenger RNA. In general, such an

-7-

antisense molecule will usually be at least 15 nucleotides, preferably, about 15-200 nucleotides, and, more preferably, 200-2,000 nucleotides in length. The antisense sequence may be complementary to all or a portion of the mRNA nucleotide sequence, and, as appreciated by those skilled in the art, the particular site or sites to which the antisense sequence binds as well as the length of the antisense sequence will vary, depending upon the degree of inhibition desired and the uniqueness of the antisense sequence. Antisense molecules may be constructed and expressed as described herein or as described, for example, in van der Krol et al., Gene 72, 45 (1988); Rodermel et al., Cell 55, 673 (1988); Mol et al., FEBS Lett. 268: 427 (1990); Weigel and Nilsson, Nature 377, 495 (1995); Cheung et al., Cell 82, 383 (1995); and U.S. Pat. No. 5,107,065.

By "obtained from" is meant isolated from or having the sequence of a naturally-occurring sequence (e.g., a cDNA, genomic DNA, synthetic DNA, or combination thereof).

By "expression control region" is meant any minimal sequence sufficient to direct transcription in a host cell (e.g., a plant or animal cell). Included in the invention are promoter elements that are sufficient to render promoter-dependent gene expression controllable for cell-, tissue-, or organ-specific gene expression, or elements that are inducible by external signals or agents (for example, light-, pathogen-, wound-, stress-, or hormone-inducible elements or chemical inducers) or elements that are capable of cycling gene transcription; such elements may be located in the 5' or 3' regions of the native gene or engineered into a transgene construct.

By "exon" is meant any length of nucleic acid sequence, excluding intron sequences and excluding the non-coding regulatory sequences driving transcription.

The exon sequence may be obtained in whole or in part from any source known in the art, including a plant, a fungus, an animal, a bacterial genome or episome, eukaryotic, nuclear or plasmid DNA, cDNA, viral DNA, or chemically synthesized DNA. If desired, an exon sequence of a target gene may contain one or more modifications in either the coding or the untranslated regions which affect the biological activity or the

-8-

chemical structure of the expression product, the rate of expression, or the manner of expression control. Such modifications include, but are not limited to, mutations, insertions, deletions, and substitutions of one or more nucleotides.

By "structural gene" is meant an uninterrupted sequence of a nucleic acid molecule, including one or more introns, bound by the appropriate splice junctions.

5 By "operably linked" is meant that a gene and a regulatory sequence(s) (e.g., an expression control region) are connected in such a way as to permit gene expression when the appropriate molecules (for example, transcriptional activator proteins) are bound to the regulatory sequence(s).

10 By "plant cell" is meant any self-propagating cell bounded by a semi-permeable membrane and containing a plastid. Such a cell also requires a cell wall if further propagation is desired. Plant cell, as used herein includes, without limitation, algae, cyanobacteria, seeds, suspension cultures, embryos, meristematic regions, callus tissue, leaves, roots, shoots, gametophytes, sporophytes, pollen, and  
15 microspores.

By "crucifer" is meant any plant that is classified within the Cruciferae family. The Cruciferae include many agricultural crops, including, without limitation, rape (for example, *Brassica campestris* and *Brassica napus*), broccoli, cabbage, brussel sprouts, radish, kale, Chinese kale, kohlrabi, cauliflower, turnip, rutabaga,  
20 mustard, horseradish, and *Arabidopsis*.

By "transgene" is meant any piece of DNA which is inserted by artifice into a cell, and becomes part of the genome of the organism which develops from that cell. Such a transgene may include a gene which is partly or entirely heterologous (i.e., foreign) to the transgenic organism, or may represent a gene homologous to an  
25 endogenous gene of the organism.

By "transgenic" is meant any cell which includes a nucleic acid molecule (e.g., a recombinant DNA molecule) which is inserted by artifice into a cell and becomes part of the genome of the organism which develops from that cell. As used herein, the transgenic organisms are generally transgenic plants or non-human

-9-

animals, and the DNA (transgene) is inserted by artifice into a nuclear or plastidic genome.

By "introducing into an animal or plant a nucleic acid" includes any number of ways of introducing a gene sequence into a eukaryotic cell that include, 5 without limitation, retroviruses, transfection, liposomes, *Agrobacterium*-mediated transformation, and biolistics, such that the introduced gene is maintained as an episome or incorporated into the genome of the cell, and is maintained and replicates as the cell divides and replicates.

By "gene therapy" is meant any therapy by which a biological deficit is 10 repaired by the introduction of a nucleic acid sequence that encodes a gene product capable of improving or delaying the disease.

By "target gene" is meant a gene that is pre-selected for gene silencing.

By "silencing" is meant that the expression of a gene is inhibited, suppressed, or decreased. Preferably, a gene is considered silenced when its 15 expression is inhibited, suppressed, or decreased, for example, by at least 50% of the normal expression level, preferably by at least 75%, and more preferably by at least 90%.

By "genetically-engineered" is meant to splice, modify, or alter a nucleic acid molecule using standard recombinant molecular techniques or methodologies. A 20 DNA molecule, for example, may be modified or altered to include site specific mutations into one or more coding and/or noncoding regions which result, upon transcription, in the production of an RNA molecule having a stem-loop structure. Such techniques may also be used to engineer a mutation, including, but not limited to an insertion, deletion, or substitution of one or more nucleotides of a nucleic acid 25 molecule (e.g., a DNA molecule).

By "agronomically important trait" is meant a trait or characteristic found in a plant that contributes to its economic value. Such traits include, without limitation, stress tolerance or resistance or both, as exemplified by resistance or tolerance to drought, heat, chilling, freezing, excessive moisture, salt stress; oxidative

-10-

stress; increased yields; food content and makeup; physical appearance; male sterility; starch properties; quantity and quality of secondary compounds; sugar quantity and quality; oil quantity and quality; and protein quantity and quality. Accordingly, one may desire to silence one or more genes involved in a regulatory or biosynthetic pathway as a means of conferring any such desirable agronomically important trait or traits on a plant.

The invention is useful for modifying the activity of an enzyme; to provide for the specific or preferential expression of an allele or one or members of a multigene family; or for the expression of a particular isozyme. In addition, the invention is useful for modulating metabolic pathways or plant development.

Other features and advantages of the invention will be apparent from the following description of the preferred embodiments thereof, and from the claims.

#### Detailed Description

The drawing will first be described.

15 Drawing

Figure 1 shows exemplary mechanisms for utilizing dsRNA-based gene silencing for the synthesis of cRNA (complementary RNA). The cRNA is typically not capped and does not have include a polyA<sup>+</sup> tail, which enhances the degradation of the mRNA encoded by a preselected target gene, and for its transport through phloem and plasmodesmata of plants.

Figure 2 is a photograph showing the effects of overexpressing sense and antisense *AtHXX1* genes in transgenic *Arabidopsis* seedlings germinated on 7% glucose/MS. The transgenic seedling expressing both sense and antisense constructs is shown on the left; the seedling expressing sense *AtHXX1* is shown on the right.

25 Gene Silencing Expression Vectors

The present invention includes novel nucleic acid expression vectors which provide a general means for silencing the expression of pre-selected target

-11-

genes. The expression vectors of the invention generally include an expression control region and an isolated nucleic acid molecule (e.g., a DNA molecule) having substantial sequence identity to a gene whose expression is selected to be silenced (e.g., an endogenous gene). Upon transcription, the DNA molecule generates an RNA 5 molecule. A portion of this generated RNA molecule then folds into a double-stranded (ds) RNA stem-loop or a region of dsRNA. Also included in the invention is a method of silencing gene expression utilizing the expression vector of the invention.

Nucleic acid molecules utilized in the invention can be obtained by several methods. Sequences for specific genes or stem-loop structures, for example, can be 10 obtained from published sequences and can also be found in the GenBank computer database (Center for Biotechnology Information; <http://www.ncbi.nlm.nih.gov/>), or may be determined after isolation using standard techniques. Nucleic acids can then be chemically synthesized, if desired, by standard methods.

Double-Stranded RNA Stem-Loops

15 As is discussed above, the expression vector of the invention which is useful for silencing the expression of an endogenous gene or virtually any other targeted gene includes a DNA molecule that, upon transcription, generates an RNA molecule having a stem-loop structure. Preferably, the stem-loop structure, within the generated RNA molecule, is positioned distally relative to the region of the RNA 20 molecule encoding the sequence of the gene selected for silencing. While not wanting to be bound by a particular theory, it is believed that the stem-loop structure provides a binding site or template for a dsRNA-dependent RNA polymerase. dsRNA-dependent RNA polymerase in turn synthesizes a cRNA based on the RNA template of the targeted gene. Exemplary strategies for designing expression constructs that 25 express RNA molecules useful for the synthesis of cRNA are shown in Fig. 1. The synthesis of the cRNA is thought to promote RNA degradation and concomitant silencing of the target gene.

The stem-loop structure refers to a nucleic acid structure that folds preferably into a hairpin form. Stem-loop structures used in the expression vectors of

-12-

the invention are typically unmodified, naturally-occurring structures, and are readily engineered for incorporation into the expression vector using standard cloning techniques. Alternatively, standard methods may be employed to synthesize stem-loop structures that mimic the naturally-occurring structures. Such genetically 5 engineered stem-loop structures, like naturally-occurring stem-loop structures, can then be integrated into the expression vector of the invention. It will be understood that the folding pattern of the stem-loop structure is not compromised by alterations in the nucleic acid sequence of the naturally-occurring molecule. For example, it is understood that alterations which include, but are not limited to, mutations, insertions, 10 deletions, and substitutions of one or more nucleotides can be made within the sequence of the stem-loop, as long as the hairpin formation of the stem-loop is maintained. Hairpin structures useful in the invention are preferably 100-200 base pairs or more in length.

In addition, a DNA molecule may be genetically engineered to express an 15 RNA molecule that forms a dsRNA structure based on complementarity found between nucleotides in its 5' and 3' regions (Fig. 1). Preferably, the dsRNA structure is approximately 20-50 bp, or is of a size sufficient to provide a template for a dsRNA-dependent RNA polymerase. The remaining non-complementary region of the RNA molecule has substantial identity to a sequence encoding a preselected target 20 gene.

dsRNA-dependent RNA polymerase

Constructs expressing a dsRNA-dependent RNA polymerase may also be introduced in a cell according to standard methods known in the art. Exemplary polymerases are described in Schiebel et al. (Plant Cell 10: 2087, 1998) and Cogoni et 25 al. (Nature 399: 166, 1999). Expression of a dsRNA-dependent RNA polymerase facilitates RNA degradation and, consequently, gene silencing. For example, if a cell does not express an endogenous dsRNA-dependent RNA polymerase or expresses an insufficient level of the enzyme, a heterologous dsRNA-dependent RNA polymerase

-13-

from any number of organisms (e.g., a plant) may be introduced to promote the generation of sequence-specific cRNAs.

Target Sequences

The expression vector of the invention also includes an isolated nucleic acid molecule (e.g., a DNA molecule) that encodes an RNA molecule corresponding to a pre-selected target gene (e.g., an endogenous gene of a cell) whose expression is to be silenced. Thus, the nucleic acid molecule once expressed in the cell silences the expression of virtually any target gene. For example, if hexokinase gene expression is to be silenced in a cell, then the expression vector of the invention incorporates a DNA molecule that is substantially identical to an endogenous hexokinase gene of the cell.

The DNA molecule utilized in the expression vector of the invention is generally obtained directly from the target gene that is to be silenced and differs, as is discussed above, primarily by the inclusion of a region which, upon transcription, folds into a dsRNA stem-loop structure. Furthermore, the DNA molecule, if desired, may be chemically synthesized. The DNA molecule encoding the RNA molecule has generally over 90% identity to the target gene to be silenced. However, DNA molecules having less than 90% identity to the target gene sequence may also be utilized; for example, a DNA molecule encoding an RNA molecule will also be effective when it shares at least 80% sequence identity with the target gene to be silenced.

While a DNA molecule may encode an RNA molecule corresponding to the full length sequence of the target gene, substantially less than the full length of the target gene can be used. For example, a DNA molecule corresponding to a portion of the target gene may also be utilized. Accordingly, relatively short regions of the target gene may be used to engineer a DNA molecule depending on the size of the target gene to be silenced.

In particular examples, in some applications, the DNA molecule of the expression vector may be 15-200 nucleotides in length, although longer sequences,

-14-

such as 100-250 nucleotides may also be utilized. In such applications where such shorter sequences (e.g., less than 200 nucleotides) are utilized, the introduced DNA molecule will share a high degree of identity with the target gene sequence, being preferably at least 80% or even 90% identical. In other applications, a sequence of 5 greater than 500 nucleotides may be used, again depending on the size of the target gene to be silenced.

Thus, the DNA molecule employed in the expression vector of the invention generally includes one contiguous nucleic acid sequence of at least 15-20 nucleotides in length that has at least 80% identity to a sequence found in the target 10 gene to be silenced. In other applications, the DNA molecule includes at least one contiguous nucleic acid sequence of at least 100 nucleotides in length that has at least 80% identity to the target gene. And, in still other applications, the DNA molecule includes at least one contiguous nucleic acid sequence of at least 250 nucleotides in length that has at least 85% identity to the target gene.

Because a full length sequence is unnecessary, the invention is not limited to the silencing of individual target genes. It is also possible to silence the expression of multiple genes using a single introduced DNA molecule that contains sequences corresponding to several genes (e.g., members of a multigene family) that are expressed in the target cell. Alternatively, it is possible to silence the expression of 20 specific members of a multigene family using a single introduced DNA molecule that contains a gene-specific sequence (e.g., sequences typically found in the 3' untranslated regions of a mRNA transcript) corresponding to the specific gene family member to be silenced in the target cell.

Once a target gene to be silenced is identified, a DNA molecule having 25 substantial identity to the target gene sequence is manipulated to render the RNA molecule encoded by the DNA molecule to include a stem-loop structure, using standard cloning methodologies. Furthermore, the proximal region of the DNA molecule itself may be tailored and manipulated using standard *in vitro* mutagenesis (e.g., using those methods described below) to modify the molecule so that upon

-15-

transcription stable dsRNA stem-loop structure formation is facilitated.

Additionally, the RNA molecule encoded by the DNA molecule may be rendered untranslatable by employing standard methods. For example, the translation initiation codon (e.g., an ATG) can be removed from the DNA sequence.

- 5     Alternatively, additional translational initiation codons may be introduced into the DNA molecule. Furthermore, the reading frame of the DNA sequence can be displaced by the addition of one or more bases. Stop codons may also be introduced after initiation codons. Many standard DNA mutagenesis techniques are available to introduce such modifications into the DNA molecule including, for example, M13 primer mutagenesis. Details of such techniques are provided in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, New York. (1989), and in Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, NY (1998).

Expression Control Regions

- 15    A wide range of expression control regions may be utilized in the expression vector of the invention, including viral, plant, or mammalian promoters. Additionally, cell- or tissue-specific promoters can, if desired, be used to trigger expression of a nucleic acid molecule in specific cell populations.

- 20    Mammalian and viral promoters suitable for use in the present invention are available in the art. The choice of a particular promoter as an expression control region of the invention, is dictated by the spatial and temporal pattern of expression that is desired for the expression of the target gene encoded by the DNA molecule of the expression vector, and ultimately for the location where the target gene is to be silenced. Expression control regions of the invention therefore include eukaryotically derived promoters which predominantly direct expression in, for example, the reproductive system (e.g., breast, ovary, testes); the musculoskeletal system (e.g., muscle or joint tissue); the cardiovascular system (e.g., capillaries or heart); the respiratory system (e.g., lung or nasal passages); the urological system (e.g., kidney or bladder); the gastrointestinal system (e.g., pancreas, liver, or intestines); the immune
- 25

-16-

system (e.g., thymus, spleen, or circulating immunological cells); the endocrine system (e.g., pituitary, gonads, and thyroid); the nervous system (e.g., neurons); or the hematopoietic system (e.g., bone marrow and peripheral blood). Further, promoters of the invention include, but are not limited to, the elastase promoter (expression in 5 pancreatic acinar cells); the alpha-A-crystallin promoter (expression in the eye lens tissue); the insulin promoter (expression in the pancreatic beta cells), and the albumin promoter regions.

Additionally, non-eukaryotically-derived promoters such as virally-derived and prokaryotically-derived promoters are included in the present invention. Such 10 promoters include, but are not limited, to the mouse mammary tumor virus promoter (MMTV), SV40 early region promoter, Rous sarcoma virus (RSV) promoter, or cytomegalovirus (CMV) promoter, which direct expression of viral or host genes in specific tissues and in many cell types.

In addition to the above-described expression control regions, a number of 15 gene promoters which are active in plant cells have been described in the literature. An example of a useful plant promoter according to the invention is a caulimovirus promoter, for example, a cauliflower mosaic virus (CaMV) promoter. These promoters confer high levels of expression in most plant tissues, and the activity of these promoters is not dependent on virally encoded proteins. CaMV is a source for 20 both the 35S and 19S promoters. In most tissues of transgenic plants, the CaMV 35S promoter is a strong promoter (see, e.g., Odell et al., Nature 313:810, 1985). The CaMV promoter is also highly active in monocots (see, e.g., Dekeyser et al., Plant Cell 2:591, 1990; Terada and Shimamoto, Mol. Gen. Genet. 220:389, 1990). Moreover, activity of this promoter can be further increased (i.e., between 2-10 fold) 25 by duplication of the CaMV 35S promoter (see e.g., Kay et al., Science 236:1299, 1987; Ow et al., Proc. Natl. Acad. Sci., U.S.A. 84:4870, 1987; and Fang et al., Plant Cell 1:141, 1989). In addition, a minimal 35S promoter may also be used as is described herein.

Other useful plant promoters include, without limitation, the nopaline

-17-

synthase promoter (An et al., *Plant Physiol.* 88:547, 1988) and the octopine synthase promoter (Fromm et al., *Plant Cell* 1:977, 1989). Exemplary monocot promoters include, without limitation, commelina yellow mottle promoter, sugar cane badna virus promoter, rice tungro baciliform virus promoter, maize streak virus element, and 5 wheat dwarf virus promoter.

For certain applications in plants, it may be desirable to express the DNA molecule of the expression construct in an appropriate tissue, at an appropriate level, or at an appropriate developmental time. For these purposes, there are an assortment of gene promoters, each with its own distinct characteristics embodied in its 10 regulatory sequences, shown to be regulated in response to the environment, hormones, and/or developmental cues. These include gene promoters that are responsible for heat-regulated gene expression (see, e.g., Callis et al., *Plant Physiol.* 88:965, 1988; Takahashi and Komeda, *Mol. Gen. Genet.* 219:365, 1989; and 15 Takahashi et al., *Plant J.* 2:751, 1992), light-regulated gene expression (e.g., the *Arabidopsis Cab2* photosynthetic, leaf specific promoter described by Mitra et al., *Plant Mol. Biol.* 12: 169-179, 1989; the pea *rbcS-3A* described by Kuhlemeier et al., *Plant Cell* 1:471, 1989; the maize *rbcS* promoter described by Schäffner and Sheen, *Plant Cell* 3:997, 1991; or the chlorophyll a/b-binding protein gene found in pea 20 described by Simpson et al., *EMBO J.* 4:2723, 1985), hormone-regulated gene expression (for example, the abscisic acid (ABA) responsive sequences from the *E<sub>m</sub>* gene of wheat described by Marcotte et al., *Plant Cell* 1:969, 1989; the ABA-inducible HVA1 and HVA22, and rd29A promoters described for barley and 25 *Arabidopsis* by Straub et al., *Plant Cell* 6:617, 1994, Shen et al., *Plant Cell* 7:295, 1995; and wound-induced gene expression (for example, of *wun1* described by Siebertz et al., *Plant Cell* 1:961, 1989), organ-specific gene expression (for example, of the tuber-specific storage protein gene described by Roshal et al., *EMBO J.* 6:1155, 1987; the 23-kDa zein gene from maize described by Schernthaner et al., *EMBO J.* 7:1249, 1988; or the French bean β-phaseolin gene described by Bustos et al., *Plant Cell* 1:839, 1989; the vegetative storage protein promoter (soybean *vspB*) described

-18-

by Sadka et al (Plant Cell 6: 737-749, 1994)), cycling promoters (e.g., the *Arabidopsis* cdc2a promoter described by Hemerly et al., Proc Natl Acad Sci USA 89: 3295-3299, 1992), senescence-specific promoters (e.g., the *Arabidopsis* SAG12 promoter described by Gan et al, Science: 270, 1986-1988, 1995), seed-specific promoters (for example, endosperm-specific or embryo-specific promoters), meristem-specific promoters (for example, the *Tch4* promoter described by Xu et al., Plant Cell 7:1555-67, 1995), or pathogen-inducible promoters (for example, PR-1 or  $\beta$ -1,3 glucanase promoters).

Transformation

Upon construction of the expression vector of the invention, several standard methods are available for introduction of the vector into a host cell (e.g., a plant or an animal). The method of transformation is not critical to the invention. Any method which provides for efficient transformation may be employed. As newer methods are available to transform host cells, they may be directly applied.

15 Gene Silencing By Overexpression of Sense and Antisense RNAs

Gene silencing is also accomplished by the simultaneous overexpression of sense and antisense RNA transcripts having substantial identity to a gene targeted for regulation. The invention therefore includes a method of silencing a gene by introducing into a cell a sense DNA construct of the targeted gene and its antisense counterpart. Using this simple approach, virtually any gene (e.g., an endogenous gene) can be regulated by the methods of the invention.

Expression vectors designed to overexpress sense and antisense transcripts of a gene targeted for regulation are constructed according to standard methods (e.g., those described herein). For example, a eukaryotic expression vector system (e.g., a cell- or tissue-specific vector) may be used to express a DNA molecule (in the sense or antisense orientation on either the same or individual vectors). Appropriate gene sequences are introduced into a plasmid or other vector, which construct is then used to transform living cells according to standard methods. Constructs containing the

-19-

entire open reading frame, inserted as a translatable or untranslatable transcript, may be used. Alternatively, portions of a DNA molecule, may be inserted. One of skill in the art will readily be able to determine which sequences to use as the appropriate sense and antisense nucleic acid construct.

5       The gene silencing effect on a targeted gene using the methods of the invention is typically monitored using standard methods. For example, the effectiveness of the simultaneous overexpression of sense and antisense genes in silencing gene expression can be determined phenotypically or by standard Northern blot analysis or immunohistochemically. Other standard nucleic acid detection  
10 techniques or alternatively immunodiagnostic techniques known to those of skill in the art (e.g., Western or Northwestern blot analysis) may be employed. Typically, a gene is considered silenced when its expression is suppressed or inhibited for example, by at least 50% of the normal expression, preferably by at least 75%, and more preferably by at least 90%.

15       The following example is provided for the purpose of illustrating the invention, and should not be construed as limiting.

      To demonstrate the application of the invention, a transgenic plant was generated which overexpressed both a sense transcript and an antisense transcript (~ 1,500 base pairs (bp)) complementary to the sense transcript. Using the hexokinase  
20 (HXK) *AtHXK1* gene of the crucifer *Arabidopsis*, sense and antisense expression constructs were generated as described in Sheen et al., WO97/35965. Gene constructs overexpressing either the sense or antisense RNAs of the coding sequences from the HXK1 gene were introduced into separate *Arabidopsis* lines. These lines were then crossed to produce plants overexpressing both sense and antisense RNAs. Plants  
25 resulting from the crosses displayed a stronger glucose insensitive phenotype than control plants. The HXK mutant was easily identified based on its ability to develop roots and expanded green cotyledons (Fig. 2). Additionally, these new HXK mutants were observed to display a stronger glucose insensitive phenotype than the null *AtHXK1* mutant or the antisense *AtHXK1* transgenic plant. The glucose insensitive

-20-

phenotype was taken as an indication that the expression of the HXK gene was silenced or knocked out by formation of double-stranded (ds) HXK RNA and resultant silencing of the endogenous *Arabidopsis* HXK gene. Finally, although plants overexpressing sense *AtHXK1* were found to have 20-50 times more AtHXK1 protein, no AtHXK1 protein was observed in plants overexpressing both the sense and antisense *AtHXK1* constructs. Similar results could be obtained by introducing sense and antisense DNA constructs in one expression vector, and introducing such an expression vector into a plant using standard methods.

In another example of simultaneously expressing sense and antisense constructs in the same plant, transgenic *Arabidopsis* plants expressing a full-length CDPK1 (calcium dependent protein kinase 1) cDNA were first generated as described in Sheen, WO 98/26045, using kanamycin as a selectable marker. Plants overexpressing the CDPK1 cDNA were observed as drought resistant. A homozygous sense CDPK1 T3 line was then generated, and two different antisense CDPK1 constructs (~ 250 bp and 700 bp), that shared identity between the CDPK1 and CDPK1a genes, were separately introduced into this line using standard *Agrobacterium* mediated plant cell transformation. Transgenic plants expressing both the sense and antisense constructs were then identified using bar and kanamycin selection. In plants expressing both sense and antisense constructs, the drought resistance phenotype was eliminated, and these plants were further found to be more sensitive to drought conditions than wild type plants. In addition, plants expressing both sense and antisense constructs were observed to have late flowering and aerial leaf phenotypes.

25    Uses

The invention described herein is useful for a variety of agricultural, horticultural, and medicinal purposes. For example, the expression vector(s) and methods described above are useful for silencing the expression of virtually any targeted gene or a particular metabolic pathway.

-21-

In the agricultural and horticultural fields, for example, the invention is particularly useful for silencing the expression of a variety of genes including, without limitation, genes associated with, for example, agronomically important traits, such as the synthesis or metabolism of peptides, proteins, fatty acids, lipids, waxes, oils,

5 starches, sugars, carbohydrates, flavors, odors, fragrances, toxins, carotenoid pigments, hormones, cell wall polymers, gene regulatory molecules, flavonoids, storage proteins, phenolic acids, coumarins, alkaloids, quinones, lignins, glucosinolates, tannins, aliphatic amines, celluloses, polysaccharides, glycoproteins, and glycolipids.

10 For instance, an alteration in the production of fatty acids or lipids can be engineered (and fatty acid composition of, e.g., an oil-producing plant thus altered) by blocking synthesis of a specific chain elongation or desaturation enzyme. Also, the synthesis of starch and sugars can be reduced (and sugar content of, e.g., an edible plant thus altered) by blocking enzymes required for starch and carbohydrate synthesis. In addition, ethylene production can be blocked to delay senescence. Similarly, fragrant molecules can be released from cells (thus altering the scent characteristics of, e.g., ornamental flowers) by blocking the enzymes responsible for glycosylation of such molecules. In addition, male-sterile plants may be engineered by silencing genes responsible for the formation of the male gametes.

15 20 In medicine, the invention provides a method of gene therapy for the treatment of cell proliferative or immunologic disorders and diseases such as those that are mediated by various overexpressed proteins. Exemplary cell-proliferative disorders include malignant as well as non-malignant cell populations which often appear to differ from the surrounding tissue both morphologically and genotypically.

25 Such disorders may be associated with abnormal expression of a gene, for example, an increased level of expression, as well as expression of a mutant form of a gene, such that the normal function of the gene product is altered. Abnormal expression also includes inappropriate expression during the cell cycle or in an incorrect cell type. The above-described methods and vectors of the invention are useful in treating

-22-

such malignancies of the various organ systems. For example, such therapy would achieve its therapeutic effect following introduction of the expression vector of the invention into cells having the proliferative disorder. Delivery of the vector can be achieved using any standard recombinant expression vectors. Various viral vectors 5 which can be utilized for gene therapy include, without limitation, adenovirus, adeno-associated virus, herpes virus, vaccinia, or an RNA virus such as a retrovirus.

The expression vectors of the invention are also useful in treating malignancies of the various organ systems, such as, for example, lung, breast, lymphoid, gastrointestinal, and genito-urinary tract as well as adenocarcinomas which 10 include malignancies such as most colon cancers, renal-cell carcinoma, prostate cancer, leukemia, breast cancer, non-small cell carcinoma of the lung, cancer of the small intestine, and cancer of the esophagus.

The methods and vectors of the invention are also useful in gene replacement therapy to reconstitute expression of a wild-type protein in cells 15 expressing a defective protein or, in the case of a heterozygote, a defective and a wild type protein. Therefore, the invention provides a method of silencing gene expression as described above and further includes administering a modified nucleic acid encoding a wild-type polypeptide corresponding to the gene product of the gene being suppressed. The replacement gene would then provide a protein with the 20 correct amino acid sequence utilizing codons with nucleotides that are not recognized by the sequences encoded by expression vectors of the invention. It is therefore possible both to silence the expression of a defective copy of a gene and to introduce a normal copy unaffected by the silencing expression vector.

There are also a number of inherited diseases in which defective genes 25 may be replaced. These include, without limitation, lysosomal storage diseases, such as those involving glucocerebrosidase, deficiencies in hypoxanthine phosphoribosyl transferase activity (the "Lesch-Nyhan" syndrome), and amyloid polyneuropathies (prealbumin).

In addition, in mammals, pathologic disorders that can be targeted using the methods

-23-

and vectors of the invention include, but are not limited, to viral infections, inflammatory disorders, cardiovascular disease, cancers, genetic disorders, and autoimmune diseases.

The methods of the present invention are also useful for suppression of a dominant negative mutation. The term "dominant negative" refers to expression of a gene resulting in a gene product that actively interferes with the function of a wild type endogenous gene product (e.g., a protein). Thus, a mutant protein or dominantly active

gene product, such as amyloid precursor protein that accumulates in Alzheimer's disease can be blocked. The dominant negative phenotype is conveyed by the expression of the mutant protein that interferes with the function of the normal protein. Such an effect is similar to the effect of dominance of one allele of a pair of alleles encoding homologous genes such that the phenotypic effect of the one allele exerts a deleterious controlling influence over the other allele.

Transgenic animals (e.g., transgenic non-human mammals) and plants can also be developed using the novel expression vectors and methods of the invention to identify the impact of increased or decreased gene expression on a particular pathway or phenotype. Standard protocols useful in producing such transgenic animals and plants are known in the art. Such protocols, for example, generally follow conventional techniques for introduction of expressible transgenes into animals such as mammals. Those of ordinary skill in the art will be familiar with these applications and will be readily able to apply the techniques in the context of the present invention. Accordingly, the expression vectors of the invention may be used to introduce DNA sequences into the germ line cells of non-humans to create transgenic animals. The preferred animal of the invention is a mouse. However, other non-humans of the invention include, but are not limited to, other rodents (e.g. rat, hamster), rabbits, chickens, sheep, goats, fish, pigs, cattle, and non-human primates.

Similarly, those skilled in the art will be familiar with methods for introducing transgenes into plants and will be able to apply the techniques in the

-24-

context of the present invention. Accordingly, the expression vectors of the invention may be used to introduce DNA sequences into plant cells, and plant cells transformed with the plant expression vector can be regenerated, for example, from single cells, callus tissue, or leaf discs according to standard plant tissue culture techniques. Exemplary plants which are useful for generating the transgenic plants (or plant cells, plant tissues, plant organs, or plant parts) of the invention, include, without limitation, dicots and monocots, such as sugar cane, wheat, rice, maize, sugar beet, barley, grape, manioc, crucifer, mustard, potato, soybean, sorghum, cassava, banana, grape, oats, tomato, millet, coconut, orange, rye, cabbage, apple, eggplant, watermelon, canola, cotton, carrot, pepper, strawberry, peanut, legume, bean, pea, mango, and sunflower.

All publications and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each independent publication or patent application was specifically and individually indicated to be incorporated by reference.

What is claimed is:

-25-

Claims

1. A method for silencing the expression of an endogenous gene in a cell, said method comprising overexpressing in said cell an isolated nucleic acid molecule of said endogenous gene and an antisense molecule comprising a nucleic acid molecule complementary to said nucleic acid molecule of said endogenous gene,  
5 wherein the overexpression of said nucleic acid molecule of said endogenous gene and said antisense molecule in said cell silences the expression of said endogenous gene.
2. The method of claim 1, wherein said isolated nucleic acid molecule of  
10 said endogenous gene generates an untranslatable RNA molecule.
3. The method of claim 1, wherein said cell is an animal cell, a mammalian cell, a cancer cell, or a plant cell.
4. The method of claim 1, wherein said isolated nucleic acid molecule of said endogenous gene comprises an exon.
- 15 5. The method of claim 4, wherein said exon is between 20-2,000 base pairs.
6. The method of claim 4, wherein overexpression of said exon and said antisense molecule silences the expression of a gene family or a specific member of a gene family.
- 20 7. The method of claim 1, wherein said endogenous gene encodes a metabolic enzyme, a structural protein, or a gene product that is associated with a disorder.

-26-

8. The method of claim 7, wherein said disorder comprises cancer, tissue inflammation, or a dysfunction in a metabolic pathway.

9. The method of claim 1, wherein said endogenous gene encodes a gene product that when silenced confers on a plant an agronomically important trait.

5 10. The method of claim 1, wherein expression of said exon or antisense molecule is regulated by a tissue-specific or cell-specific expression control region or by a constitutive expression control region.

11. A method for silencing the expression of an endogenous gene in a cell, said method comprising expressing in said cell a vector comprising

10 (i) an expression control region functional in said cell; and  
(ii) an operably linked DNA molecule comprising a proximal region and a distal region, said proximal region having substantial sequence identity to an endogenous gene of said cell;

15 wherein transcription of said DNA generates an RNA molecule having a genetically-engineered dsRNA stem-loop structure based on complementarity between nucleotides found in said distal region of said RNA molecule.

12. The method of claim 11, wherein said expression vector is overexpressed in said cell.

13. The method of claim 11, wherein said generated RNA molecule is 20 untranslatable.

14. The method of claim 11, wherein said cell is an animal cell, a mammalian cell, a cancer cell, or a plant cell.

-27-

15. The method of claim 11, wherein the expression of said DNA molecule silences the expression of a gene family or a specific member of a gene family.

16. The method of claim 11, wherein said endogenous gene encodes a  
5 metabolic enzyme, a structural protein, or a gene product that is associated with a disorder.

17. The method of claim 16, wherein said disorder comprises cancer, tissue inflammation, or a dysfunction in a metabolic pathway.

18. The method of claim 11, wherein said endogenous gene encodes a  
10 gene product that when silenced confers on a plant an agronomically important trait.

19. The method of claim 11, wherein of said expression control region is a tissue-specific or cell-specific expression control region or a constitutive expression control region.

20. The method of claim 11, wherein said DNA comprises an exon.

15 21. The method of claim 20, wherein said exon is between 20-2,000 base pairs.

22. A gene silencing expression vector comprising  
(i) an expression control region functional in said cell; and  
(ii) an operably linked DNA molecule comprising a proximal region and  
20 a distal region, said proximal region having substantial sequence identity  
to an endogenous gene of said cell;  
wherein transcription of said DNA generates an RNA molecule having a genetically-

-28-

engineered dsRNA stem-loop structure based on complementarity between nucleotides found in said distal region of said RNA molecule.

23. The expression vector of claim 22, wherein said generated RNA molecule is untranslatable.

5 24. The expression vector of claim 22, wherein the expression of said DNA molecule silences the expression of a gene family or a specific member of a gene family.

10 25. The expression vector of claim 22, wherein said expression control region is a tissue-specific or cell-specific expression control region or is a constitutive expression control region.

26. The expression vector of claim 22, wherein said DNA comprises an exon.

27. The expression vector of claim 26, wherein said exon is between 20-2,000 base pairs.

15 28. A method for silencing the expression of an endogenous gene in a cell, said method comprising expressing in said cell a vector comprising  
(i) an expression control region functional in said cell; and  
(ii) an operably linked DNA molecule having substantial sequence identity to an endogenous gene of said cell; wherein transcription of said DNA generates an RNA molecule having a dsRNA structure based on complementarity between nucleotides found in the 5' and 3' regions of said RNA molecule.

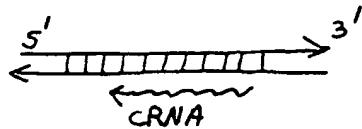
-29-

29. A gene silencing expression vector comprising  
(i) an expression control region functional in said cell; and  
(ii) an operably linked DNA molecule having substantial sequence  
identity to an endogenous gene of said cell; wherein transcription of said DNA  
generates an RNA molecule having a dsRNA based on complementarity between  
nucleotides found in the 5' and 3' regions of said RNA molecule.

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70  
75  
80  
85  
90  
95  
100  
105  
110  
115  
120  
125  
130  
135  
140  
145  
150  
155  
160  
165  
170  
175  
180  
185  
190  
195  
200  
205  
210  
215  
220  
225  
230  
235  
240  
245  
250  
255  
260  
265  
270  
275  
280  
285  
290  
295  
300  
305  
310  
315  
320  
325  
330  
335  
340  
345  
350  
355  
360  
365  
370  
375  
380  
385  
390  
395  
400  
405  
410  
415  
420  
425  
430  
435  
440  
445  
450  
455  
460  
465  
470  
475  
480  
485  
490  
495  
500  
505  
510  
515  
520  
525  
530  
535  
540  
545  
550  
555  
560  
565  
570  
575  
580  
585  
590  
595  
600  
605  
610  
615  
620  
625  
630  
635  
640  
645  
650  
655  
660  
665  
670  
675  
680  
685  
690  
695  
700  
705  
710  
715  
720  
725  
730  
735  
740  
745  
750  
755  
760  
765  
770  
775  
780  
785  
790  
795  
800  
805  
810  
815  
820  
825  
830  
835  
840  
845  
850  
855  
860  
865  
870  
875  
880  
885  
890  
895  
900  
905  
910  
915  
920  
925  
930  
935  
940  
945  
950  
955  
960  
965  
970  
975  
980  
985  
990  
995  
1000  
1005  
1010  
1015  
1020  
1025  
1030  
1035  
1040  
1045  
1050  
1055  
1060  
1065  
1070  
1075  
1080  
1085  
1090  
1095  
1100  
1105  
1110  
1115  
1120  
1125  
1130  
1135  
1140  
1145  
1150  
1155  
1160  
1165  
1170  
1175  
1180  
1185  
1190  
1195  
1200  
1205  
1210  
1215  
1220  
1225  
1230  
1235  
1240  
1245  
1250  
1255  
1260  
1265  
1270  
1275  
1280  
1285  
1290  
1295  
1300  
1305  
1310  
1315  
1320  
1325  
1330  
1335  
1340  
1345  
1350  
1355  
1360  
1365  
1370  
1375  
1380  
1385  
1390  
1395  
1400  
1405  
1410  
1415  
1420  
1425  
1430  
1435  
1440  
1445  
1450  
1455  
1460  
1465  
1470  
1475  
1480  
1485  
1490  
1495  
1500  
1505  
1510  
1515  
1520  
1525  
1530  
1535  
1540  
1545  
1550  
1555  
1560  
1565  
1570  
1575  
1580  
1585  
1590  
1595  
1600  
1605  
1610  
1615  
1620  
1625  
1630  
1635  
1640  
1645  
1650  
1655  
1660  
1665  
1670  
1675  
1680  
1685  
1690  
1695  
1700  
1705  
1710  
1715  
1720  
1725  
1730  
1735  
1740  
1745  
1750  
1755  
1760  
1765  
1770  
1775  
1780  
1785  
1790  
1795  
1800  
1805  
1810  
1815  
1820  
1825  
1830  
1835  
1840  
1845  
1850  
1855  
1860  
1865  
1870  
1875  
1880  
1885  
1890  
1895  
1900  
1905  
1910  
1915  
1920  
1925  
1930  
1935  
1940  
1945  
1950  
1955  
1960  
1965  
1970  
1975  
1980  
1985  
1990  
1995  
2000  
2005  
2010  
2015  
2020  
2025  
2030  
2035  
2040  
2045  
2050  
2055  
2060  
2065  
2070  
2075  
2080  
2085  
2090  
2095  
2100  
2105  
2110  
2115  
2120  
2125  
2130  
2135  
2140  
2145  
2150  
2155  
2160  
2165  
2170  
2175  
2180  
2185  
2190  
2195  
2200  
2205  
2210  
2215  
2220  
2225  
2230  
2235  
2240  
2245  
2250  
2255  
2260  
2265  
2270  
2275  
2280  
2285  
2290  
2295  
2300  
2305  
2310  
2315  
2320  
2325  
2330  
2335  
2340  
2345  
2350  
2355  
2360  
2365  
2370  
2375  
2380  
2385  
2390  
2395  
2400  
2405  
2410  
2415  
2420  
2425  
2430  
2435  
2440  
2445  
2450  
2455  
2460  
2465  
2470  
2475  
2480  
2485  
2490  
2495  
2500  
2505  
2510  
2515  
2520  
2525  
2530  
2535  
2540  
2545  
2550  
2555  
2560  
2565  
2570  
2575  
2580  
2585  
2590  
2595  
2600  
2605  
2610  
2615  
2620  
2625  
2630  
2635  
2640  
2645  
2650  
2655  
2660  
2665  
2670  
2675  
2680  
2685  
2690  
2695  
2700  
2705  
2710  
2715  
2720  
2725  
2730  
2735  
2740  
2745  
2750  
2755  
2760  
2765  
2770  
2775  
2780  
2785  
2790  
2795  
2800  
2805  
2810  
2815  
2820  
2825  
2830  
2835  
2840  
2845  
2850  
2855  
2860  
2865  
2870  
2875  
2880  
2885  
2890  
2895  
2900  
2905  
2910  
2915  
2920  
2925  
2930  
2935  
2940  
2945  
2950  
2955  
2960  
2965  
2970  
2975  
2980  
2985  
2990  
2995  
3000  
3005  
3010  
3015  
3020  
3025  
3030  
3035  
3040  
3045  
3050  
3055  
3060  
3065  
3070  
3075  
3080  
3085  
3090  
3095  
3100  
3105  
3110  
3115  
3120  
3125  
3130  
3135  
3140  
3145  
3150  
3155  
3160  
3165  
3170  
3175  
3180  
3185  
3190  
3195  
3200  
3205  
3210  
3215  
3220  
3225  
3230  
3235  
3240  
3245  
3250  
3255  
3260  
3265  
3270  
3275  
3280  
3285  
3290  
3295  
3300  
3305  
3310  
3315  
3320  
3325  
3330  
3335  
3340  
3345  
3350  
3355  
3360  
3365  
3370  
3375  
3380  
3385  
3390  
3395  
3400  
3405  
3410  
3415  
3420  
3425  
3430  
3435  
3440  
3445  
3450  
3455  
3460  
3465  
3470  
3475  
3480  
3485  
3490  
3495  
3500  
3505  
3510  
3515  
3520  
3525  
3530  
3535  
3540  
3545  
3550  
3555  
3560  
3565  
3570  
3575  
3580  
3585  
3590  
3595  
3600  
3605  
3610  
3615  
3620  
3625  
3630  
3635  
3640  
3645  
3650  
3655  
3660  
3665  
3670  
3675  
3680  
3685  
3690  
3695  
3700  
3705  
3710  
3715  
3720  
3725  
3730  
3735  
3740  
3745  
3750  
3755  
3760  
3765  
3770  
3775  
3780  
3785  
3790  
3795  
3800  
3805  
3810  
3815  
3820  
3825  
3830  
3835  
3840  
3845  
3850  
3855  
3860  
3865  
3870  
3875  
3880  
3885  
3890  
3895  
3900  
3905  
3910  
3915  
3920  
3925  
3930  
3935  
3940  
3945  
3950  
3955  
3960  
3965  
3970  
3975  
3980  
3985  
3990  
3995  
4000  
4005  
4010  
4015  
4020  
4025  
4030  
4035  
4040  
4045  
4050  
4055  
4060  
4065  
4070  
4075  
4080  
4085  
4090  
4095  
4100  
4105  
4110  
4115  
4120  
4125  
4130  
4135  
4140  
4145  
4150  
4155  
4160  
4165  
4170  
4175  
4180  
4185  
4190  
4195  
4200  
4205  
4210  
4215  
4220  
4225  
4230  
4235  
4240  
4245  
4250  
4255  
4260  
4265  
4270  
4275  
4280  
4285  
4290  
4295  
4300  
4305  
4310  
4315  
4320  
4325  
4330  
4335  
4340  
4345  
4350  
4355  
4360  
4365  
4370  
4375  
4380  
4385  
4390  
4395  
4400  
4405  
4410  
4415  
4420  
4425  
4430  
4435  
4440  
4445  
4450  
4455  
4460  
4465  
4470  
4475  
4480  
4485  
4490  
4495  
4500  
4505  
4510  
4515  
4520  
4525  
4530  
4535  
4540  
4545  
4550  
4555  
4560  
4565  
4570  
4575  
4580  
4585  
4590  
4595  
4600  
4605  
4610  
4615  
4620  
4625  
4630  
4635  
4640  
4645  
4650  
4655  
4660  
4665  
4670  
4675  
4680  
4685  
4690  
4695  
4700  
4705  
4710  
4715  
4720  
4725  
4730  
4735  
4740  
4745  
4750  
4755  
4760  
4765  
4770  
4775  
4780  
4785  
4790  
4795  
4800  
4805  
4810  
4815  
4820  
4825  
4830  
4835  
4840  
4845  
4850  
4855  
4860  
4865  
4870  
4875  
4880  
4885  
4890  
4895  
4900  
4905  
4910  
4915  
4920  
4925  
4930  
4935  
4940  
4945  
4950  
4955  
4960  
4965  
4970  
4975  
4980  
4985  
4990  
4995  
5000  
5005  
5010  
5015  
5020  
5025  
5030  
5035  
5040  
5045  
5050  
5055  
5060  
5065  
5070  
5075  
5080  
5085  
5090  
5095  
5100  
5105  
5110  
5115  
5120  
5125  
5130  
5135  
5140  
5145  
5150  
5155  
5160  
5165  
5170  
5175  
5180  
5185  
5190  
5195  
5200  
5205  
5210  
5215  
5220  
5225  
5230  
5235  
5240  
5245  
5250  
5255  
5260  
5265  
5270  
5275  
5280  
5285  
5290  
5295  
5300  
5305  
5310  
5315  
5320  
5325  
5330  
5335  
5340  
5345  
5350  
5355  
5360  
5365  
5370  
5375  
5380  
5385  
5390  
5395  
5400  
5405  
5410  
5415  
5420  
5425  
5430  
5435  
5440  
5445  
5450  
5455  
5460  
5465  
5470  
5475  
5480  
5485  
5490  
5495  
5500  
5505  
5510  
5515  
5520  
5525  
5530  
5535  
5540  
5545  
5550  
5555  
5560  
5565  
5570  
5575  
5580  
5585  
5590  
5595  
5600  
5605  
5610  
5615  
5620  
5625  
5630  
5635  
5640  
5645  
5650  
5655  
5660  
5665  
5670  
5675  
5680  
5685  
5690  
5695  
5700  
5705  
5710  
5715  
5720  
5725  
5730  
5735  
5740  
5745  
5750  
5755  
5760  
5765  
5770  
5775  
5780  
5785  
5790  
5795  
5800  
5805  
5810  
5815  
5820  
5825  
5830  
5835  
5840  
5845  
5850  
5855  
5860  
5865  
5870  
5875  
5880  
5885  
5890  
5895  
5900  
5905  
5910  
5915  
5920  
5925  
5930  
5935  
5940  
5945  
5950  
5955  
5960  
5965  
5970  
5975  
5980  
5985  
5990  
5995  
6000  
6005  
6010  
6015  
6020  
6025  
6030  
6035  
6040  
6045  
6050  
6055  
6060  
6065  
6070  
6075  
6080  
6085  
6090  
6095  
6100  
6105  
6110  
6115  
6120  
6125  
6130  
6135  
6140  
6145  
6150  
6155  
6160  
6165  
6170  
6175  
6180  
6185  
6190  
6195  
6200  
6205  
6210  
6215  
6220  
6225  
6230  
6235  
6240  
6245  
6250  
6255  
6260  
6265  
6270  
6275  
6280  
6285  
6290  
6295  
6300  
6305  
6310  
6315  
6320  
6325  
6330  
6335  
6340  
6345  
6350  
6355  
6360  
6365  
6370  
6375  
6380  
6385  
6390  
6395  
6400  
6405  
6410  
6415  
6420  
6425  
6430  
6435  
6440  
6445  
6450  
6455  
6460  
6465  
6470  
6475  
6480  
6485  
6490  
6495  
6500  
6505  
6510  
6515  
6520  
6525  
6530  
6535  
6540  
6545  
6550  
6555  
6560  
6565  
6570  
6575  
6580  
6585  
6590  
6595  
6600  
6605  
6610  
6615  
6620  
6625  
6630  
6635  
6640  
6645  
6650  
6655  
6660  
6665  
6670  
6675  
6680  
6685  
6690  
6695  
6700  
6705  
6710  
6715  
6720  
6725  
6730  
6735  
6740  
6745  
6750  
6755  
6760  
6765  
6770  
6775  
6780  
6785  
6790  
6795  
6800  
6805  
6810  
6815  
6820  
6825  
6830  
6835  
6840  
6845  
6850  
6855  
6860  
6865  
6870  
6875  
6880  
6885  
6890  
6895  
6900  
6905  
6910  
6915  
6920  
6925  
6930  
6935  
6940  
6945  
6950  
6955  
6960  
6965  
6970  
6975  
6980  
6985  
6990  
6995  
7000  
7005  
7010  
7015  
7020  
7025  
7030  
7035  
7040  
7045  
7050  
7055  
7060  
7065  
7070  
7075  
7080  
7085  
7090  
7095  
7100  
7105  
7110  
7115  
7120  
7125  
7130  
7135  
7140  
7145  
7150  
7155  
7160  
7165  
7170  
7175  
7180  
7185  
7190  
7195  
7200  
7205  
7210  
7215  
7220  
7225  
7230  
7235  
7240  
7245  
7250  
7255  
7260  
7265  
7270  
7275  
7280  
7285  
7290  
7295  
7300  
7305  
7310  
7315  
7320  
7325  
7330  
7335  
7340  
7345  
7350  
7355  
7360  
7365  
7370  
7375  
7380  
7385  
7390  
7395  
7400  
7405  
7410  
7415  
7420  
7425  
7430  
7435  
7440  
7445  
7450  
7455  
7460  
7465  
7470  
7475  
7480  
7485  
7490  
7495  
7500  
7505  
7510  
7515  
7520  
7525  
7530  
7535  
7540  
7545  
7550  
7555  
7560  
7565  
7570  
7575  
7580  
7585  
7590  
7595  
7600  
7605  
7610  
7615  
7620  
7625  
7630  
7635  
7640  
7645  
7650  
7655  
7660  
7665  
7670  
7675  
7680  
7685  
7690  
7695  
7700  
7705  
7710  
7715  
7720  
7725  
7730  
7735  
7740  
7745  
7750  
7755  
7760  
7765  
7770  
7775  
7780  
7785  
7790  
7795  
7800  
7805  
7810  
7815  
7820  
7825  
7830  
7835  
7840  
7845  
7850  
7855  
7860  
7865  
7870  
7875  
7880  
7885  
7890  
7895  
7900  
7905  
7910  
7915  
7920  
7925  
7930  
7935  
7940  
7945  
7950  
7955  
7960  
7965  
7970  
7975  
7980  
7985  
7990  
7995  
8000  
8005  
8010  
8015  
8020  
8025  
8030  
8035  
8040  
8045  
8050  
8055  
8060  
8065  
8070  
8075  
8080  
8085  
8090  
8095  
8100  
8105  
8110  
8115  
8120  
8125  
8130  
8135  
8140  
8145  
8150  
8155  
8160  
8165  
8170  
8175  
8180  
8185  
8190  
8195  
8200  
8205  
8210  
8215  
8220  
8225  
8230  
8235  
8240  
8245  
8250  
8255  
8260  
8265  
8270  
8275  
8280  
8285  
8290  
8295  
8300  
8305  
8310  
8315  
8320  
8325  
8330  
8335  
8340  
8345  
8350  
8355  
8360  
8365  
8370  
8375  
8380  
8385  
8390  
8395  
8400  
8405  
8410  
8415  
8420  
8425  
8430  
8435  
8440  
8445  
8450  
8455  
8460  
8465  
8470  
8475  
8480  
8485  
8490  
8495  
8500  
8505  
8510  
8515  
8520  
8525  
8530  
8535  
8540  
8545  
8550  
8555  
8560  
8565  
8570  
8575  
8580  
8585  
8590  
8595  
8600  
8605  
8610  
8615  
8620  
8625  
8630  
8635  
8640  
8645  
8650  
8655  
8660  
8665  
8670  
8675  
8680  
8685  
8690  
8695  
8700  
8705  
8710  
8715  
8720  
8725  
8730  
8735  
8740  
8745  
8750  
8755  
8760  
8765  
8770  
8775  
8780  
8785  
8790  
8795  
8800  
8805  
8810  
8815  
8820  
8825  
8830  
8835  
8840  
8845  
8850  
8855  
8860  
8865  
8870  
8875  
8880  
8885  
8890  
8895  
8900  
8905  
8910  
8915  
8920  
8925  
8930  
8935  
8940  
8945  
8950  
8955  
8960  
8965  
8970  
8975  
8980  
8985  
8990  
8995  
9000  
9005  
9010  
9015  
9020  
9025  
9030  
9035  
9040  
9045  
9050  
9055  
9060  
9065  
9070  
9075  
9080  
9085  
9090  
9095  
9100  
9105  
9110  
9115  
9120  
9125  
9130  
9135  
9140  
9145  
9150  
9155  
9160  
9165  
9170  
9175  
9180  
9185  
9190  
9195  
9200  
9205  
9210  
9215  
9220  
9225  
9230  
9235  
9240  
9245  
9250  
9255  
9260  
9265  
9270  
9275  
9280  
9285  
9290  
9295  
9300  
9305  
9310  
9315  
9320  
9325  
9330  
9335  
9340  
9345  
9350  
9355  
9360  
9365  
9370  
9375  
9380  
9385  
9390  
9395  
9400  
9405  
9410  
9415  
9420  
9425  
9430  
9435  
9440  
9445  
9450

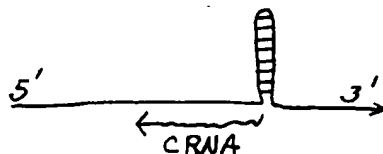
### dsRNA-based Silencing

#### Sense and antisense RNA



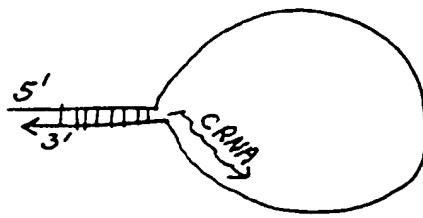
cRNA is important for gene-specific mRNA degradation and concomitant gene silencing.

#### Sense RNA with dsRNA at the 3' end



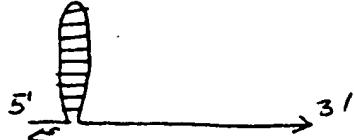
dsRNA is the target for dsRNA-dependent RNA polymerase, which synthases cRNA. The dsRNA region need not include the sequence of the gene targeted for RNA degradation and, consequently, gene silencing.

#### Sense RNA with 5' and 3' repeats to form dsRNA



The sequence of cRNA produced from the RNA template is important for gene silencing, but not the sequence of the dsRNA.

#### Sense RNA with dsRNA at the 5' end



No gene silencing will result, because there is no RNA template to generate cRNA.

**Figure 1**

# BEST AVAILABLE COPY

WO 00/49035

PCT/US00/04287

2/2



At H<sub>X</sub>K knock-out      sense AtH<sub>X</sub>K1

Figure 2

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/04287

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) C07H 21/04; A01H 5/00; C12N 15/09, 15/14, 15/00  
US CL : 435/455, 419; 536/23.1, 24.5; 800/419, 295

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/455, 419; 536/23.1, 24.5; 800/419, 295

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  
NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
WEST, MEDLINE, BIOTECHNO, SCISEARCH, BIOSIS, EMBASE, CAPLUS

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	DOUGHERTY et al. Transgenes and Gene Suppression: Telling Us Something New? Current Opinion in Cell Biology. 1995. Vol. 7, pages 399-405, see entire document.	1-32
Y	HAMADA et al. Co-Suppression of the Hdrophobin Gene HCf-1 is Correlated with Antisense RNA Biosynthesis in Cladosporium fulvum. Mol. Gen. Genet. 1998. Vol. 259, pages 630-638, see entire document.	1-32
Y,P	JENSEN et al. Cosuppression of I Transposon Activity in Drosophila by I-Containing Sense and Antisense Transgenes. Genetics. 1999. Vol. 153, pages 1767-1774, see entire document.	1-32

Further documents are listed in the continuation of Box C.  See patent family annex.

* Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*'A'	document defining the general state of the art which is not considered to be of particular relevance	
*'E'	earlier document published on or after the international filing date	"X"
*'L'	document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y"
*'O'	document referring to an oral disclosure, use, exhibition or other means	"A"
*'P'	document published prior to the international filing date but later than the priority date claimed	document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
20 APRIL 2000	21 JUN 2000

Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer SEAN MCGARRY Telephone No. (703) 308-0196
---	--

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/04287

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CARTEA et al. Comparison of Sense and Antisense Methodologies for Modifying the Fatty Acid Composition of <i>Arabidopsis thalia</i> Oilseed. <i>Plant Science</i> . 1998. Vol. 136, pages 181-194, see entire document.	1-32